

## 論文の内容の要旨

論文題目	マルチモーダル潜在的ディリクレ配分法の多層化による 知識の確率的表現 ーロボットの言語獲得・行動決定への応用ー
学 位 申 請 者	アッタミミ ムハンマド

近年、ロボットと人の共存を目指すための研究が盛んに行われている。ロボットが人と自然に暮らすためには、人の言葉を理解する必要があり、その言葉の背後にある潜在的な意味を解釈して行動しなければならない。また、コミュニケーションのために、ロボット自身の意図を言語として創出することが望まれる。旧来の人工知能の研究では、単語を単なる記号として扱い、その記号で閉じた世界の中で言語を理解する努力を続けてきた。自然言語処理・理解は、この流れを強く受けている。これに対して近年のロボティクス・人工知能研究では、いわゆる記号接地問題を基本として、言語の本質的な意味を扱い始めているが、未だに言語の理解や生成の本質的な解決には遠く及ばない。

本論文では、ロボットが経験によって得るマルチモーダル情報に基づいて多様な概念を形成し、この概念を基盤とした言語理解・生成を考えることでこの問題を解決する新たな方向性を示す。ここで、概念とはマルチモーダルな情報を分類して形成される「カテゴリ」であり、この概念を通して様々な予測をすることが「理解」とであると定義する。さらに言語は、こうした概念と結びついた音韻ラベルであり、人との自然なインタラクションの中で獲得することが可能である。つまり本論文で提案するモデルは、ロボットが日常の活動によって得ることのできる情報を基盤に概念を形成し、音韻ラベルとの結びつきや語の順番を意味する文法をボトムアップに獲得することで、言語の意味理解や生成を実現するものである。

これまで先行研究において、マルチモーダル情報を用いた物体のカテゴリ分類手法が提案されている。実際に、ロボットが経験することによって得た情報をカテゴリ分類することで、人間の感覚に近い物体概念の形成が可能であることを示している。また、形成された概念を利用して未観測情報を予測することができ、ロボットによる物体の理解が前述の定義の範囲で可能であると言える。しかし、より人間のように柔軟な理解をロボットで実現するためには、物体概念の獲得だけでは不十分であることは明らかである。なぜなら、ほとんどの物体はそれを使う人や使う人の動き、使われる場所などが関連しており、これらの情報を予測で

きない限りその物体を理解したとは言えないためである。つまり、物体概念のみならず人の動き概念や場所概念など多様な概念を学習すると同時に、それらの関係性を獲得する必要がある。このような多様な概念の獲得は、先行研究の成果である単一概念の形成を基盤として、マルチモーダル情報の階層的カテゴリ分類へと発展させることで実現することで可能であり、最終的にはこれがロボットによる「事物の真の理解の計算モデル」となることを明らかにしたい。

そこで本論文ではまず、マルチモーダル情報の階層的カテゴリ分類手法を提案する。提案手法は、マルチモーダル潜在的ディリクレ配分法 (Multimodal Latent Dirichlet Allocation: MLDA) を階層化した多層マルチモーダル潜在的ディリクレ配分法 (multilayered MLDA: mMLDA) である。下層のMLDAでは下位概念である、物体、動き、場所、人物の概念がそれぞれ形成され、上層のMLDAではこれらの概念を統合する上位概念が形成される。このモデルを用いることで例えば、下位概念としてジュースという物体概念や物を口に運ぶという動き概念、ダイニングという場所概念などが形成される。上位層ではこれらの関係性が学習され、「飲む」という行動概念が形成される。これにより、ジュースを見ることでそれを口に運ぶ「飲む」という行動や、その「飲む」という行動が「ダイニング」という場所で行なわれやすいといった未観測情報の予測を行うことが可能となる。

また、形成された多様な概念を利用し、同時に語意や文法を獲得することで、観測したシーンを文章で表現する手法を検討する。先行研究において、入力されたマルチモーダル情報に対応する単語や、単語が指す概念の推論の可能性が示されている。しかし本論文で扱う問題は、階層的な概念における語意の獲得であり、どの階層のどの概念にどの単語が結び付くかという問題を解く必要があり、この問題は、先行研究では実現されていない。本論文では、単語と概念間の相互情報量を用いることで、どの単語が本来どの概念に結びついているのかを自動的に推定する手法を提案する。これにより単語と概念の結び付きを学習することが可能であり、各単語に対応する、物体、場所、人、動作といった概念クラスの推定が可能である。従って、教示発話における概念クラスの生起順を学習することで、概念クラスの遷移確率という形で表現される確率文法を学習することができる。これによって、ロボットによる言語の意味理解や生成を実現することが可能となる。

一方、実際のコミュニケーションは、背景知識や周辺状況などといった文脈を考慮しなければ成立しない。つまり、事物に対する理解をより柔軟に行うためには、学んできた多様な概念を活用した上で、様々な文脈を考慮する必要がある。本論文では、ロボットが人と生活する上で、様々な文脈においてどのように行動決定するかを議論する。つまり、獲得した多様な概念と文脈と統合することで、適切な行動を決定する手法を提案する。これにより例えば、人が普段ソファでテレビを見ているときに、お菓子を食べながらお茶を飲んでいるということを知っていれば、人が「お菓子を持ってきて」と命令した際の音声認識に誤りが生じたとしても、その時に「ソファでテレビを見ていてお茶を飲んでいる」という文脈を用いることで、ロボットが適切に判断をして正しい行動をとることができる可能性がある。

# 論文審査の結果の要旨

学位申請者氏名      ATTAMIMI MUHAMMAD

審査委員主査      長井    隆行

委員      金子    正秀

委員      田中    一男

委員      横井    浩史

委員      内田    雅文

第1章では、まず本論文の背景と目的について述べた。この章では特に、人間の知能や人工知能における理解の研究について述べ、人間が行っているような真の意味での理解が、ロボットをはじめとする人工物では、いまだに実現されておらず、未知の環境で柔軟にタスクを実現することができない、という問題提起を行った。

第2章では、前章の議論をより具体化するために、ロボットによる掃除タスクを取り上げている。現在の知能ロボットの多くは、タスクを実現するために作り込みが行われており、そうした作り込みによるタスク実現が、具体的にどのように行われるかについて説明がなされている。ここではさらにその具体的な問題を指摘し、その問題を克服するために、ロボットが概念や概念同士の関係性といった知識をボトムアップに獲得する必要があることを述べた。また、実世界の不確実性を考慮して、確率的にこうした知識を表現することが有効であることを議論した。

第3章では、2章の議論を受けて、具体的にどのように知識を表現し、さらにそれを学習するのかについて議論した。ここでは、確率モデルである、マルチモーダル潜在的ディリクレ配分法を多層化することで、知識を確率的に表現できることを明らかにした。これは、先行研究で提案されているマルチモーダル潜在的ディリクレ配分法を階層的に結合したモデルとなっており、そうしたモデルを用いることで、動きの概念化と物体の概念化を行うと同時に、それらの共起関係を上位の概念として概念化できることを示した。この上位の概念が人の動作に対応し

ていることを，実験結果を通して明らかにした．また，動きから使用する物体の予測など，未観測な情報が予測できることを示し，これによって概念を通した事物の理解や，概念間の関係の獲得による知識表現の有効性を示した．

第4章では，第3章で述べた多層マルチモーダル潜在的ディリクレ配分法を拡張し，物体と動きだけでなく，場所や人といった概念を獲得することを検討した．さらに，拡張したモデルに対して，言語を結びつける手法を提案した．この際に問題となるのは，「各単語がどの概念に属しているのか」という結びつけを自動的に見つける必要があるという問題と，「文章を生成するための文法をどのように学習するのか」という問題の2点である．前者の問題に対して，各概念と各単語の相互情報量を計算することで，単語と概念の結びつきを発見できることを示した．後者の問題に対しては，概念の種類を統語範疇と考え，その確率的な遷移を文法と捉えることで獲得することを提案した．実際のデータから多様な概念やそれらの関係性といった知識が獲得され，さらにその概念を用いて観測したシーンを言語で表現できることを実験によって示した．

第5章では，前章までで獲得した知識をさらに時間的な関係性に発展させることを検討している．つまり，多層マルチモーダル潜在的ディリクレ配分法だけでは，共起性を手掛かりとした学習は実現できるものの，時間的な推移を表現することができないという問題がある．これは，多層マルチモーダル潜在的ディリクレ配分法を時間発展させることで解決することができる．ここでは，3章のモデルを時間的に接続することで，ある行動の後にどのような行動が行われるかを予測し，さらにその際にどのような物体が使われるかを予測できることを示した．さらに，このような予測によって，ロボットが実際の家庭環境で様々なサービスを実現できる可能性があることを明らかにした．

第6章では論文全体をまとめ，今後の課題について述べた．特に，2章で述べた掃除タスクが，現状の提案でどこまで柔軟に実現可能になったのかについて議論している．明らかに，作り込みのタスクでは実現不可能なことが実現可能となったが，まだ十分ではないことも明らかにしている．特に，タスクを実行するためには知識だけでなく，その知識を使った行動決定が必要である．このような行動決定をするためには，強化学習の枠組みを提案したモデルに統合する必要がある．しかし，このような統合は，モデルの形状から考えると自然に行える可能性があり，今後こうした点を検討することが，最も重要な課題であることを述べた．

以上のように本論文は，ロボットの知能化に大きく貢献するものである．よって本論文は，博士（工学）の学位論文として十分な価値を有するものと認める．